A Contour-Based Approach to Multisensor Image Registration

Hui Li, *Student Member, IEEE*, B. S. Manjunath, *Member, IEEE*, and Sanjit K. Mitra, *Member, IEEE*

Abstract— **Image registration is concerned with the establishment of correspondence between images of the same scene. One challenging problem in this area is the registration of multispectral/multisensor images. In general, such images have different gray level characteristics, and simple techniques such as those based on area correlations cannot be applied directly. On the other hand, contours representing region boundaries are preserved in most cases. In this paper, we present two contour-based methods which use region boundaries and other strong edges as matching primitives. The first contour matching algorithm is based on the chain-code correlation and other shape similarity criteria such as invariant moments. Closed contours and the salient segments along the open contours are matched separately. This method works well for image pairs in which the contour information is well preserved, such as the optical images from Landsat and Spot satellites. For the registration of the optical images with synthetic aperture radar (SAR) images, we propose an elastic contour matching scheme based on the active contour model. Using the contours from the optical image as the initial condition, accurate contour locations in the SAR image are obtained by applying the active contour model. Both contour matching methods are automatic and computationally quite efficient. Experimental results with various kinds of image data have verified the robustness of our algorithms, which have outperformed manual registration in terms of root mean square error at the control points.**

I. Introduction

In many image processing applications it is necessary to compare multiple images of the same scene acquired by different sensors, or images taken by the same sensor but at different times. These images may have relative translation, rotation, scale, and other geometric transformations between them. The goal of image registration is to establish the correspondence between two images and determine the geometric transformation that aligns one image with the other. Registering multisensor data enables comparison and fusion of information from different sensory modalities, which often provide complementary information about the region surveyed. For example, optical images from Landsat provide information on chemical composition, vegetation, and biological properties of the surface, while radar images from Seasat provide information on surface roughness and surface electrical properties such as the dielectric constant [1]. Thus, one of the most attractive aspects of multisensor data acquisition is the fact that the sensors are collecting very different types of information from the same scene. However, because the information from sensors of different types are inherently different, the problem of registration is inevitably more complex than registration of images from the same type of sensors. Images from different channels of the multispectral imaging systems (e.g., Landsat thematic mapper (TM), Spot and NOAA advanced very high resolution radiometer (AVHRR)) bear unique characteristics determined by the wavelengths at which they are acquired. Substantial differences exist between the images from passive sensors such as Landsat TM and active sensors such as synthetic aperture radar (SAR). Comparing a Landsat image to a Seasat SAR image, the discrepancies may include cloud cover in the Landsat image, grainy appearance of the SAR image resulting from speckle noise, and radiometric reversal in the relative brightness between them [1].

Existing image registration techniques broadly fall into two categories: the area-based and the feature-based methods [2], [3]. In the area-based method, a small window of points in the first image is statistically compared with windows of the same size in the second image. The measure of match is usually the normalized cross-correlation. The centers of the matched windows are control points which can be used to solve for the transformation parameters between the two images. One innovative method in the recent literature is the computational vision approach proposed in [4]. In this approach, image rotation is obtained by taking the difference between the estimated illuminant directions, under the assumption that the images are taken at about the same time (or under the assumption of a stationary source of illumination). Feature points in the images are detected by using a biologically motivated model based on the Gabor wavelet decomposition [5]. The windows surrounding the feature points are matched with each other using an area-based correlation in a hierarchical manner. However, area-based methods like this are not well-adapted to the problem of multisensor image registration since the gray-level characteristics of images to be matched are quite different. Feature-based methods, which extract and match the common structures from two images, have been shown to be more suitable for this task [6]–[9]. One technique uses ancillary data such as high resolution digital elevation maps (DEM) of the area to be registered [6]. Simulated multisensor images with shadowing effects are generated by illuminating the DEM from sensor imaging geometries. The actual images are registered with DEM by correlating with

Manuscript received March 21, 1993; revised November 15, 1993. This work was supported by the Lawrence Livermore National Laboratory and by a University of California MICRO grant, with matching funds from ESL Inc., Digital Instruments, Inc., Tektronix, Inc., and Rockwell International Corp. The associate editor coordinating the review of this paper and approving it for publication was Prof. Rama Chellappa.

The authors are with the Center for Information Processing Research, Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106 USA.

IEEE Log Number 9408198.

the simulated data. As a result, the multisensor images are coregistered on the common grid provided by the DEM. The authors in [6] also suggest using a segmentation scheme for multisensor region boundary extraction and the Canny edge detector for edge detection. The boundary or edge maps from Landsat, Spot and Seasat are then matched by a binary cross-correlation technique. The assumptions of this approach are that the images to be registered are coarsely aligned based on prior information and that the residual rotation error is small. The scheme proposed in [7] extracts objects from Landsat and Spot images at different scales and matches them using their structural attributes such as ellipticity, thinness, and inclination. In [8] closed boundaries are extracted by a segmentation technique and used as matching primitives; however, the good control points on the nonclosed regions are neglected. In [9] a contour-based method for registering Spot and Seasat images is proposed. A long coastal line is used as a landmark and the matching is conducted in a coarse-to-fine fashion using a scale space representation. Edge-and contour-based techniques have also been used in obtaining stereo correspondence [10]–[13], but typically they assume small displacements to constrain the search space. The structural stereo matching technique proposed in [14] is capable of matching images with large rotation and translation. However, this method is better suited for registration of industrial images rather than the natural remotely sensed images from which a symbolically represented structural scene description may not be easy to establish.

In this paper, we present two contour-based image registration algorithms: a basic contour matching scheme and an elastic contour matching scheme for optical-to-SAR image registration. The proposed schemes are automatic and computationally quite efficient. The basic contour matching scheme uses a two-threshold method to extract well-defined contours and separate them into two groups: open contours and closed contours. Chain code correlation and other shape similarity criteria such as moments are used to match the closed contours. For the open contours, salient segments such as corners are detected first and then used in the matching process. A consistency check is conducted in the transformation parameter space. This first method works well for image pairs in which the contour information is well preserved, such as the optical images from Landsat and Spot satellites. Because of the coherent nature of the illumination, radar images are speckled. As a result, the usual edge detectors yield poor results when applied to radar images (e.g., SAR images) [15]. In this paper, a second method is developed for optical-to-SAR image registration. Rather than extracting contours from optical and SAR images separately and matching them later, a better strategy, as proposed here, is to use the optical contours as initial guides for contour detection in the SAR images. Assuming that the optical and the SAR images can be coarsely aligned, the active contour model provides an ideal tool for this task. Another benefit of the proposed elastic matching technique is its ability to handle deformation between the images to be matched. We have applied our algorithms on three test data sets: single sensor images, 12-band multispectral aerial images, and multisensor Landsat, Spot, and Seasat images. The proposed contour matching algorithms are robust and accurate if the boundary information is preserved through different sensory modalities and if this information can be successfully extracted. The proposed methods have outperformed manual registration in terms of root mean square error (RMSE) at the control points in our computer simulations.

This paper is organized as follows: Section II discusses the basic contour matching procedure; Section III reviews the active contour model and discusses the optical-to-SAR image registration scheme using this model; Section IV presents the experimental results; and Section V provides the concluding remarks.

II. The Contour Matching Algorithm

A. Contour Extraction

Fig. 1 shows the block diagram of the proposed basic contour matching scheme. Contour extraction is carried out in two steps. First, images $I(x, y)$ are convolved with a Laplacian-of-Gaussian (LoG) operator and the edges are detected at the zero-crossing points [2]. The LoG operator is decomposed into the sum of two separable filters to speed up computation [16]. Simple zero-crossing patterns such as "$+ + - -$" and "$- - + +$" (composed of signs of pixel values of the filtered image) are detected along both vertical and horizontal directions. In the second step, the edge strength at each zero-crossing point is computed by considering the slopes of the LoG of the image along both x and y directions (denoted by k_x and k_y). An edge strength map is defined as follows:

$$S(x, y) = \begin{cases} \sqrt{k_x^2 + k_y^2}, & \text{if } (x, y) \text{ is a zero crossing point,} \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

The contours that are retained for further processing satisfy the following conditions [17]:

i) The edge strength at each point along the contour is greater than T_1

ii) At least one point on the contour has an edge strength greater than T_2

where T_1 and T_2 are preset thresholds and $T_1 < T_2$. T_1 is set to preserve the whole contour around the region boundary without incurring discontinuities at weak edge points. T_2 is chosen large enough to avoid spurious edges. This two-threshold scheme is implemented by scanning the 2-D edge strength array. Contour search is initiated whenever one point with a value greater than T_2 is scanned. The search is conducted in both directions of the contour and the neighboring pixels with values greater than T_1 are accepted as contour points until no neighboring pixels are found satisfying this condition. Then all the edge strength values along the detected contour are set to zero such that these points will not be visited again. The same search operation continues until the whole edge strength array has been scanned. The contours are then divided into two categories—closed contours and open contours—and are represented by a chain code. Short contours that cannot be used reliably in the matching process are discarded at this point.

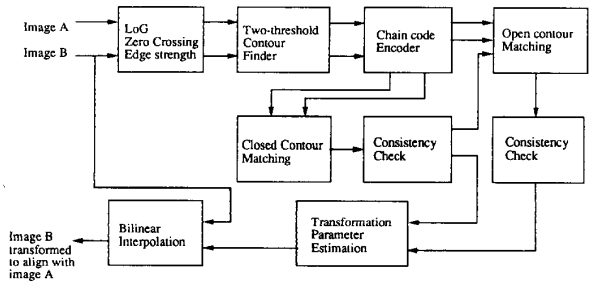

Fig. 1. Schematic diagram for the basic contour matching method.

B. Chain Code Criteria

A digital curve can be represented by an integer sequence $\{a_i \in \{0, 1, 2, \ldots, 7\}\}$, depending on the relative position of the current edge pixel with respect to the previous edge pixel [2], [18]. One unit corresponds to an angle of 45°. Thus a chain code value of 3, for example, indicates the next pixel is on the north-west (135°) direction. The standard chain code representation has certain drawbacks. For example, a line along $-22.5°$ direction is coded as $\{707070\ldots\}$. To prevent such wraparound, we convert a length n standard chain code $\{a_1 a_2 \cdots a_n\}$ into a modified code $\{b_1 b_2 \cdots b_n\}$ by a shifting operation defined recursively by:

$$\begin{cases} b_1 = a_1 \\ b_i = q_i, q_i \text{ is an integer such that } (q_i - a_i) \bmod 8 = 0 \\ \qquad \text{and } |q_i - b_{i-1}| \text{ is minimized, } i = 2, 3, \ldots n. \end{cases} \tag{2}$$

The line along $-22.5°$ direction is then coded as $\{787878\ldots\}$. The shifted chain code is further smoothed by a five-point Gaussian filter $\{0.1, 0.2, 0.4, 0.2, 0.1\}$. A comparison of the standard chain codes and the modified chain codes for a pair of matched contours[1] is shown in Fig. 2. From here on it is assumed that the chain code has been shifted and smoothed.

Let contour A be represented by an N_A-point chain code $\{a_i\}$ and let contour B be represented by an N_B-point chain code $\{b_i\}$. A correlation measure D_{kl} between two n-point segments, one starting at index k of contour A and the other one starting at index l of contour B, is defined as:

$$D_{kl} = \frac{1}{n} \sum_{j=0}^{n-1} \cos \frac{\pi}{4} (a'_{k+j} - b'_{l+j}) \tag{3}$$

where

$$a'_{k+i} = a_{(k+i) \bmod N_A} - \frac{1}{n} \sum_{j=0}^{n-1} a_{(k+j) \bmod N_A}, 0 \le i < n, \tag{4}$$

$$b'_{l+i} = b_{(l+i) \bmod N_B} - \frac{1}{n} \sum_{j=0}^{n-1} b_{(l+j) \bmod N_B}, 0 \le i < n. \tag{5}$$

Here, the modulus operation accommodates the cases in which the contours are closed. The correlation measure is similar to the mean-squared-error of two signals. The cosine function ensures $D_{kl} \le 1$, and $D_{kl} = 1$ when there is a perfect match.

[1] This pair of contours was extracted from the images in Fig. 4(a)–(b) and can be found in the contour maps in Fig. 4(c)–(d).

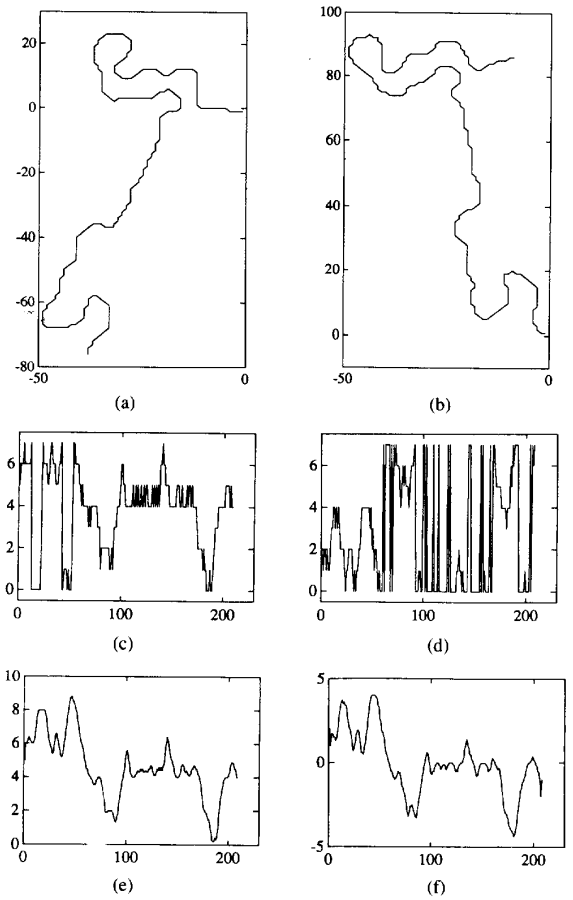

Fig. 2. Modified chain code representation: (a), (b) A pair of contours extracted from images shown in Fig. 4(a)–(b); (c), (d) the corresponding standard chain code representations; (e), (f) the modified chain codes after shifting and smoothing.

In order to identify the location of best fit between contours A and B, an n-point segment from contour A starting at index k is slid over contour B. A similarity function C_{AB} is then defined as:

$$C_{AB} = \max\{D_{kl}\}_{l \in M} \tag{6}$$

where M specifies the search range. For a pair of closed contours A and B, the whole contours can be used for the matching purpose. Suppose that $k = 0$, $n = N_A$, and that M includes every index of contour B. Then the similarity function between contours A and B becomes:

$$C_{AB} = \max\{D_{0l}\}_{0 \le l < N_B}. \tag{7}$$

The rotation between the contours is reflected in the difference between the average values of the corresponding chain code representations (see Fig. 2(e)–(f)). Since D_{kl} is normalized with respect to the mean value, the similarity function C_{AB} is invariant to rotation if the quantization effect of the chain code is neglected. The derivative of the chain code, obtained

by using the first difference, is also rotation invariant. However, we found that using the mean-subtracted chain code is more accurate in locating the positions of best matches. For the closed contour the similarity function can be further normalized with respect to contour length, and becomes scale invariant. In implementation, the chain code of the longer closed contour is simply resampled by linear interpolation to have the same number of points as the shorter closed contour. Subsequently, the similarity criterion C_{AB} is computed based on the resampled chain codes of the same length.

C. Contour Matching

The contour matching process begins with the matching of closed contours. For every closed contour, five shape attributes are computed: the perimeter, the longest and shortest distances from boundary to the centroid, and the first and second invariant moments proposed by Hu [2]. These moments were defined originally for 2-D images. For the case of 2-D contours, the first and second moments can be defined as follows:

$$h_1 = \frac{1}{n^2}\sum_{i=1}^{n}[(x_i - x_c)^2 + (y_i - y_c)^2] \tag{8}$$

$$h_2 = \frac{1}{n^4}\left[\sum_{i=1}^{n}(x_i - x_c)^2 - \sum_{i=1}^{n}(y_i - y_c)^2\right]^2 + \frac{4}{n^4}\left[\sum_{i=1}^{n}(x_i - x_c)(y_i - y_c)\right]^2. \tag{9}$$

Here x_i and y_i are the coordinates of each point along the contour, x_c and y_c are the coordinates of the centroids of the contour, and n is the length of the contour. Each closed contour in the first image is compared with every closed contours in the second image. The pairs are accepted as candidate matches if the relative differences of their shape attributes are below some preset thresholds (e.g., 20%). This step narrows down the prospective matches to a few contours with very similar shapes. The corresponding pair is then determined by the chain code based similarity function C_{AB}. Contour A in the first image and contour B in the second image are selected as a matched pair if the following two conditions are satisfied:

1) $C_{AB} \geq C_{AB'}$, where B' includes all the contours with similar shapes to contour A, and
2) $C_{AB} > T_3$, where T_3 is a preset threshold which eliminates matches with poor correlation.

In the rare case that multiple contours get matched to the same contour, the pair with highest value of C_{AB} is selected. The centroids of these matched contours are used as control points from which an initial estimation of the transformation parameters is computed to guide the second stage of open contour matching.

For matching the open contours, we use salient segments along the contours rather than the whole contours as matching primitives. Salient segments such as corners can be detected from the chain code representation. For a contour of length n with shifted and smoothed chain code $\{a_i\}$, we define a measure of curvature at the ith point as:

$$c_i = \max_{1<j<3\sigma}\{\max\{|a_{i-j} - a_{i+j}|, |a_{i-j} - a_{i+j-1}|\}\} \tag{10}$$

where σ is the standard deviation of the LoG operator used for the contour extraction. The ith point along the contour is chosen as a salient point if both of the following conditions are satisfied:

i) $c_i \geq T_4$, and
ii) $c_i \geq c_k$ for all $k \in [i - p, i + p]$.

Here p is a constant that determines the minimum distance between the salient points and T_4 is a threshold specifying the minimum acceptable curvature. For example, if $T_4 = 2$, then the salient feature locations correspond to corners where the curve bends by at least 90°. The contour segments surrounding the salient points are then used as 1-D templates in finding the corresponding matches in the other image.

D. Estimation of Transformation Parameters

It is often sufficient to assume that the relationship between a point (X, Y) in one image and its corresponding point $(\hat{X}, \hat{Y})$ in the other image can be expressed by a 2-D affine transform [4]:

$$\begin{pmatrix}\hat{X}\\ \hat{Y}\end{pmatrix} = \kappa\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}X\\ Y\end{pmatrix} + \begin{pmatrix}\Delta X\\ \Delta Y\end{pmatrix} \tag{11}$$

where parameters $(\kappa, \theta, \Delta X, \Delta Y)$ correspond to a scaling factor, a rotation angle, and translations along the two orthogonal directions, respectively. The transform parameters are estimated from a set of matched points $\{(X_i, Y_i)\}$ and $\{(\hat{X}_i, \hat{Y}_i)\}$. The problem can be solved through linearization of the above equations [4]. Suppose κ and θ can be estimated based on prior knowledge by κ_0 and θ_0, and that $\theta = \theta_0 + \Delta\theta$. Then the following two approximations can be established: $\kappa\cos\theta \approx \kappa_0\cos\theta_0 - (\kappa_0\sin\theta_0)\Delta\theta$, and $\kappa\sin\theta \approx \kappa_0\sin\theta_0 + (\kappa_0\cos\theta_0)\Delta\theta$. Next, the unknowns $(\Delta\theta, \Delta X, \Delta Y)$ can be solved by the least squares method. Another method for solving (11), as adopted in our implementation, uses variable substitutions $u = \kappa\cos\theta$ and $v = \kappa\sin\theta$ to obtain the following new equations:

$$\begin{pmatrix}\hat{X}\\ \hat{Y}\end{pmatrix} = \begin{pmatrix}u & v\\ -v & u\end{pmatrix}\begin{pmatrix}X\\ Y\end{pmatrix} + \begin{pmatrix}\Delta X\\ \Delta Y\end{pmatrix}. \tag{12}$$

Unknowns u, v, ΔX, and ΔY can be solved in the least squares sense based on all the matched points. Scaling factor κ and rotation θ can then be determined from u and v: $\kappa = \sqrt{u^2 + v^2}$, and $\theta = \arctan(\frac{v}{u})$ for $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$. More general transformations such as bivariate polynomial transformations are more effective if nonlinear distortion exists between the images [19], [20].

The root mean square error (RMSE) between the matched points after the transformation provides a measure of registration accuracy and is defined as:

$$\mathrm{RMSE} = \left(\sum_{i=1}^{m}[(uX_i + vY_i + \Delta X - \hat{X}_i)^2 + (uY_i - vX_i + \Delta Y - \hat{Y}_i)^2]/m\right)^{\frac{1}{2}} \tag{13}$$

where m is the number of matched points. It is important to mention that m must be reasonably large to make the accuracy measure meaningful (note that RMSE $= 0$ when $m = 2$, whether or not the matched points are correct).

E. A Consistency Check

In the contour matching process some false matches are inevitable. Therefore, a global consistency check is necessary to ensure correct registration. We exploit the fact that distances are preserved under a rigid transformation. Our consistency check method is closely related to the Hough transform technique. Let $\overline{A_1A_2}$ denote the distance between points A_1 and A_2. For two sets of m matched points $\{A_i\}$ and $\{B_i\}$, ratios of $\overline{A_iA_j}/\overline{B_iB_j}$ are computed based on all $\frac{m(m-1)}{2}$ possible combinations. The resulting histogram of scale should form a cluster corresponding to the true scale difference between the images. The pairs that contribute to the cluster will be accepted as correct matches while the pairs whose contribution is scattered and away from the peak are declared as mismatches and discarded. The actual consistency check is done in an iterative fashion. The most likely mismatches are rejected first, followed by the computation of RMSE based on the remaining matched points. If RMSE is too large, another round of consistency check is carried out. The iteration continues until either an acceptable RMSE is achieved or the number of retained matches is less than three, which means no matched points can be found. Since RMSE is a good indication whether a set of correct matched points has been found, this scheme is capable of eliminating most false matches.

III. Optical-to-SAR Image Registration Using an Active Contour Model

The presence of speckle noise due to the coherent nature of the illumination makes it difficult to detect good edge and contour information in SAR images [9], [15], [21]. Various smoothing techniques [22], [23] can be applied before edge detection, but this in turn may affect the accuracy of the contour location. Similar performance limitations are characteristic of statistical edge operators [1]. Good results based on the Canny edge detector were reported in [6]; however, many contours corresponding to continuous boundaries were broken. In order to match such contours with contours from an optical image, sophisticated heuristics will be needed to perform edge linking for the contours from the SAR image. The LoG edge detector can produce unbroken contours, but also generates false contours because of the textural patterns present in the SAR images. Consequently, for optical-to-SAR image registration the basic contour matching method presented earlier can only work in limited cases and produces much fewer control points than optical-to-optical image registration.

In contrast to SAR images, optical images have fairly well defined edges and hence are easier to work with. Here, we suggest the use of edge information provided by optical images to better localize the contours in the SAR images. It was shown in [6] and [24] that multisensor data such as Landsat, Spot, and Seasat images can be rectified and geocoded to a common earth grid and resampled to the same pixel spacing within 100 meters or 5 pixels. Based on this assumption, the contours from the optical images can be directly superimposed over SAR images as prior information. The second contour-matching algorithm proposed in this paper involves the use of an active contour model [25], [26] for optical-to-SAR image registration. We first briefly review the idea of the active contour model and discuss our modification, and then apply it to the problem of image registration.

The active contour model is an elastic contour extraction technique that makes use of prior information about the shape of the region boundary to obtain a close fit to the data. The use of prior information helps in robust detection of contours while minimizing the undesirable effects of noise and texture in the image data. The elastic contour extraction is usually formulated as an optimization problem involving an energy functional and an initial contour which specifies the region of interest. The energy functional can be decomposed into three parts: internal, image, and external energy. The internal energy regulates the smoothness and deformation of the contour while the image energy attracts points on the contour toward the true edge locations in the image. The external energy influences the contour deformation through feedback from other sources of information. In recent years various elastic contour extraction methods have been reported, based on variational calculus [25], dynamic programming [27], and a local descent method (greedy algorithm) [26].

We review here the greedy approach described in [26] which considers only the internal and the image forces. Let a contour be represented by a sequence of vectors: $\mathbf{v}(s) = (x(s), y(s))$, where $x(s)$ and $y(s)$ are the pixel coordinates. The index range is $1 \leq s \leq N$, where N is the number of points. Let $I(x, y)$ be the image function. An energy functional corresponding to the contour is defined as:

$$E = \sum_{s=1}^{N} \alpha(s)E_{\text{cont}} + \beta(s)E_{\text{curv}} - \gamma(s)E_{\text{image}} \quad (14)$$

$$E_{\text{cont}} = |\mathbf{v}_s - \mathbf{v}_{s-1}| - \frac{1}{N-1}\left(\sum_{i=2}^{N} |\mathbf{v}_i - \mathbf{v}_{i-1}|\right) \quad (15)$$

$$E_{\text{curv}} = |\mathbf{v}_{s-1} - 2\mathbf{v}_s + \mathbf{v}_{s+1}|^2 \quad (16)$$

$$E_{\text{image}} = |\bigtriangledown I(x(s), y(s))|^2. \quad (17)$$

E_{cont} and E_{curv} represent the internal energy of the contour. E_{cont} maintains a uniform distribution of points on the contour while E_{curv} enforces a smoothness constraint. The image force E_{image} is the gradient magnitude. $\alpha(s)$, $\beta(s)$, and $\gamma(s)$ are the weighting parameters. Starting with the initial contour locations, the neighborhood of each point on the contour is examined to obtain a potential new location for that point. The contour point is moved to the new location if it results in a lower energy. The iterations continue until a minimum energy is obtained.

Since edge detectors such as the LoG operator or the Canny detector usually include a Gaussian smoothing operation to determine the scale of the edges to be extracted, it is appropriate to also incorporate this feature into the image force. Considering the speckle noise in the SAR images, we chose a geometric filter [22] combined with a LoG operator for the

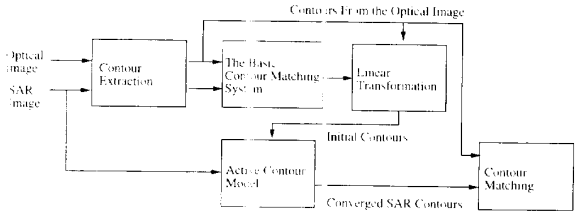

Fig. 3. Optical-to-SAR image registration using the active contour model.

image force. The geometric filter works as follows: For any three consecutive pixels along the horizontal direction, the center pixel is decreased by 1 if its intensity value is larger than its two neighbors; whereas, it is increased by 1 if its intensity value is smaller than its neighbors. The same operations are also applied to any three consecutive pixels along the vertical and two diagonal directions. Multiple iterations of the above operations are usually required. The modified image force is defined by a 2-D convolution as follows:

$$E_{\text{image}} = S(x, y) * G(x, y) \tag{18}$$

where $G(x, y)$ is a 2-D Gaussian function and $S(x, y)$ is the edge strength map as defined in (1). The modified image force provides a larger bowl of attraction, hence enabling convergence to a better local minimum of the corresponding energy functional.

For the problem of optical-to-SAR image registration, the contours extracted from the optical image can be used as the initial condition for the contour detection of the SAR image. In this way the active contour model can be implemented without any human intervention. As mentioned earlier, multisensor images can be roughly aligned (in many cases) within 5 pixels based on the knowledge of the sensor platforms. Thus, the contours from the optical image can be directly superimposed over the SAR image.

If the knowledge of the sensor platforms is not available and the rectified multisensor data are not initially aligned well, the optical and SAR images have to be preregistered manually or using some other automatic algorithms. In our experiments we have used either the basic contour matching algorithm presented earlier or the Fourier transform-based technique presented in [28] to perform the preregistration. The schematic diagram of the overall image registration system in which the basic contour matching scheme is adopted as the first stage matching is shown in Fig. 3. In this case a heavily smoothed SAR image is used instead of the original one for the contour extraction.

When the optical image has a higher resolution than the SAR image, quite often some optical contours have no corresponding region boundaries in the SAR image. To discard mistakes caused by the spurious initial contour, a converged contour produced by the active contour model algorithm is accepted only if the average image force ($\frac{1}{N}\sum_{s=1}^{N} E_{\text{image}}$) is above a preset threshold. This step helps to ensure that enough contour points lie over or close to the true edge locations, as indicated by the image force. At this point, the correspondence between contours in the two images is known (but not point-to-point correspondence). From this set of corresponding contours it is easy to obtain a set of control points. For the closed contours the centroids are used as control points. For the open contours salient segments such as corners are detected in one contour and are matched with the other contour using the similarity function defined in (6). The resulting set of control points is subject to a consistency check and is used to compute the least square error estimates of the transformation parameters.

TABLE I
COMPARISON OF CONTOUR-BASED AUTOMATIC REGISTRATION RESULTS WITH THE MANUAL REGISTRATION RESULTS OR WITH GROUND TRUTH. THE UNIT FOR ΔX, ΔY AND RMSE IS IN PIXELS.

Test data	Method	Matches	Scale	$\Delta\theta$	ΔX	ΔY	RMSE
optical	manual	31	1.005	−29.95°	275.82	−189.39	1.82
images	contour matching	42	1.0045	−29.95°	275.74	−188.21	0.69
12 bend	ground truth		1.0000	15.00°	7.00	250.00	
FM	manual	25	0.9999	−15.23°	6.51	249.50	1.68
simulator	contour matching	35	1.0010	−15.02°	7.38	249.63	0.61
Landsat	manual	28	1.9993	1.07°	0.15	0.42	1.50
Spot (1)	contour matching	28	0.9997	1.15°	0.19	0.70	0.81
Landsat	manual	27	1.0001	−88.93°	0.37	510.83	1.51
Spot (2)	contour matching	27	0.9991	−88.81°	1.03	510.82	0.87
Landsat	manual	7	1.3336	−15.86°	36.15	173.61	1.20
Spot (3)	contour matching	7	1.3303	−15.89°	37.60	173.56	1.00
Landsat	manual	25	0.9966	10.60°	27.29	68.31	2.56
Seasat	active contour	25	1.0001	10.53°	29.09	67.90	2.11
Spot	manual	27	1.0011	−0.12°	1.07	3.04	1.72
Seasat	active contour	27	1.0006	0.56°	1.75	3.07	3.11

IV. EXPERIMENTAL RESULTS

In this section, we first provide experimental results for single sensor image registration, multispectral image registration, and multisensor optical image registration using the basic contour matching scheme. Although our emphasis is on multisensor registration, we include a single sensor example to illustrate the robustness of the algorithm for images without distinct features. We next show results of optical-to-SAR (Landsat-to-Seasat and Spot-to-Seasat) image registration based on the active contour model. The standard deviation of the LoG operator is 3 for all examples. The thresholds for the two-threshold contour extraction scheme are fine-tuned to each type of image. The typical values are $T_1 = 10$ and $T_2 = 100$, when the maximum edge strength is normalized to 255. The parameters for the corner detection are set to $T_4 = 1.8$ and $p = 25$. The length of salient segments used in open contour matching is 31. The weighting parameters for the active contour model are constants: $\alpha(s) = 0.1$, $\beta(s) = 0.8$, and $\gamma(s) = 1.2$. To evaluate the performance of our algorithms, we manually selected distinct landmarks as control points. A comparison of transformation parameters and RMSE obtained by manual registration and contour-based automatic registration is shown in Table I. The number of manually selected control points and the number of control points generated by the proposed algorithms are the same to ensure a fair comparison of RMSE. Our algorithms are reasonably efficient in terms of computational complexity. The size of the images in all examples is 512 × 512. The contour extraction and contour matching take about 90 to 100 s on

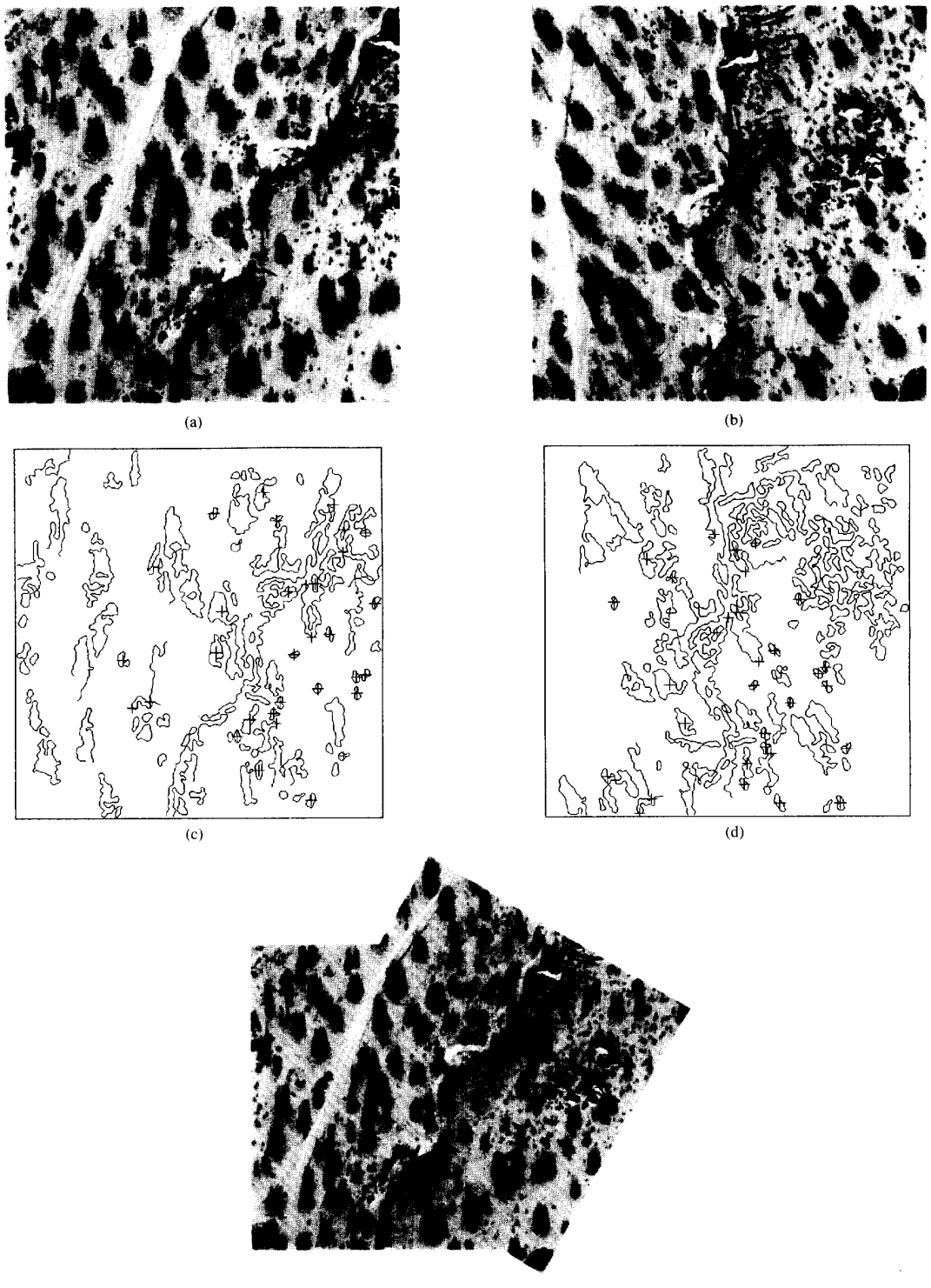

Fig. 4. (a), (b) Two images from a Mojave Desert sequence; (c), (d) the matched contours with the "+" denoting the centroids of the closed contours and the salient points along the open contours; (e) mosaic of (a) and (b). (Test data courtesy of JPL, Caltech.)

a Sun Sparc2 workstation. The active contour model used in optical-to-SAR image registration takes about 20 to 40 s.

Single-Sensor Optical Image Registration: Figs. 4(a) and (b) show two images from a Mojave Desert sequence taken

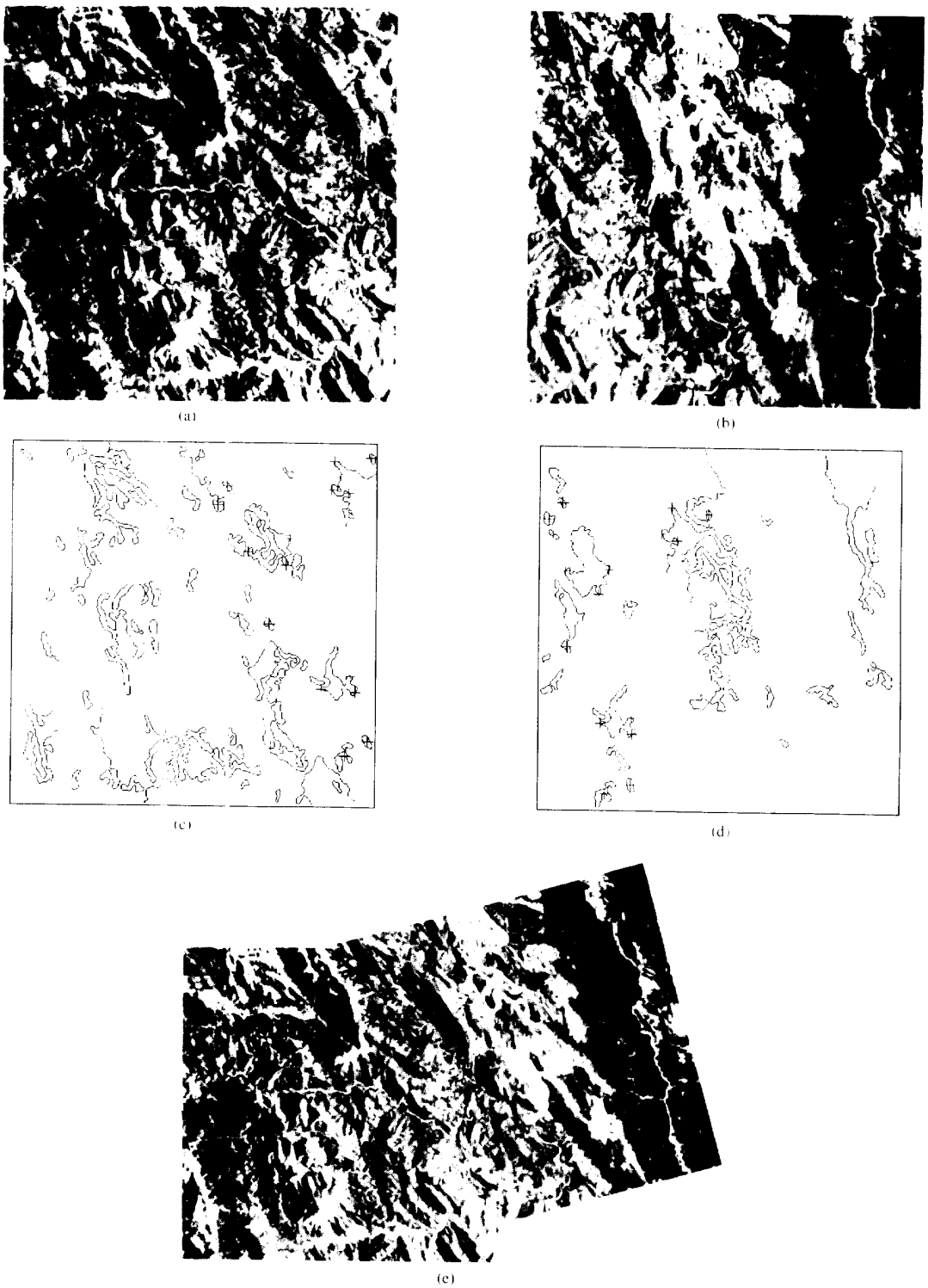

Fig. 5. (a), (b) Zeroth and eighth bands, respectively, of 12-band aerial image data; (c), (d) matched contours with the "+" marks denoting the centroids of the closed contours and the salient points along the open contours; (e) mosaic of (a) and (b).

with an optical camera. Twenty-four pairs of closed contours and seven pairs of salient points along open contours were matched. Fig. 4(c)–(d) show the matched contours, with "+" denoting the centroids of closed contours and the matched

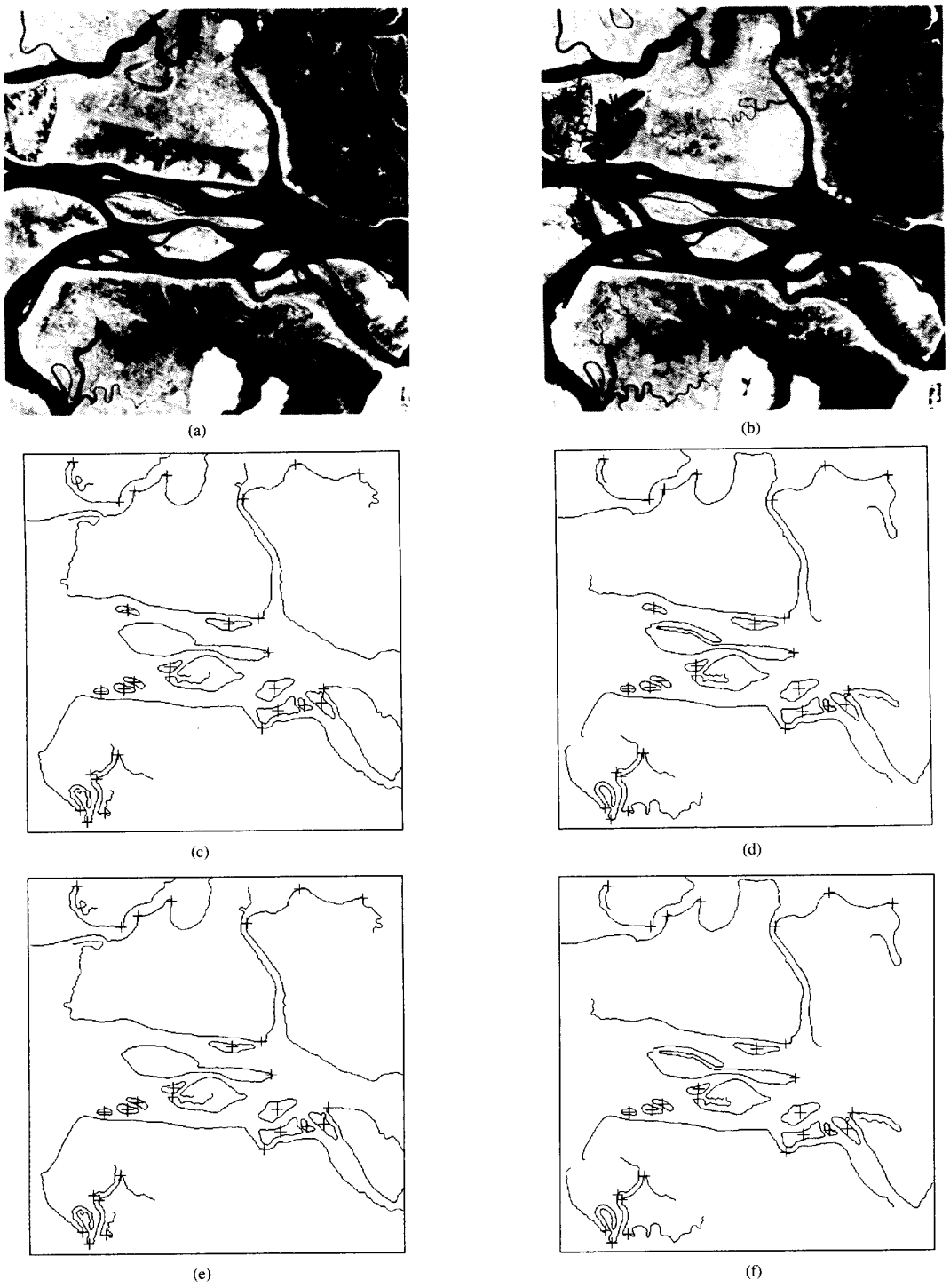

Fig. 6. (a), (b) Landsat image and a Spot image, respectively; (c), (d) matched contours of (a) and (b); (e), (f) matched contours of the Landsat image and the 90°-rotated Spot image.

segments of open contours. Fig. 4(e) shows the mosaic of (a) and (b).

Multispectral Optical Image Registration: A set of 12-band aerial images of a mountainous area acquired by a thematic mapper (TM) simulator was used in this experiment. The size of each raw image is 1000 × 766 and the images are perfectly registered with each other. Test images (512 X 512) were cut out from different bands of this

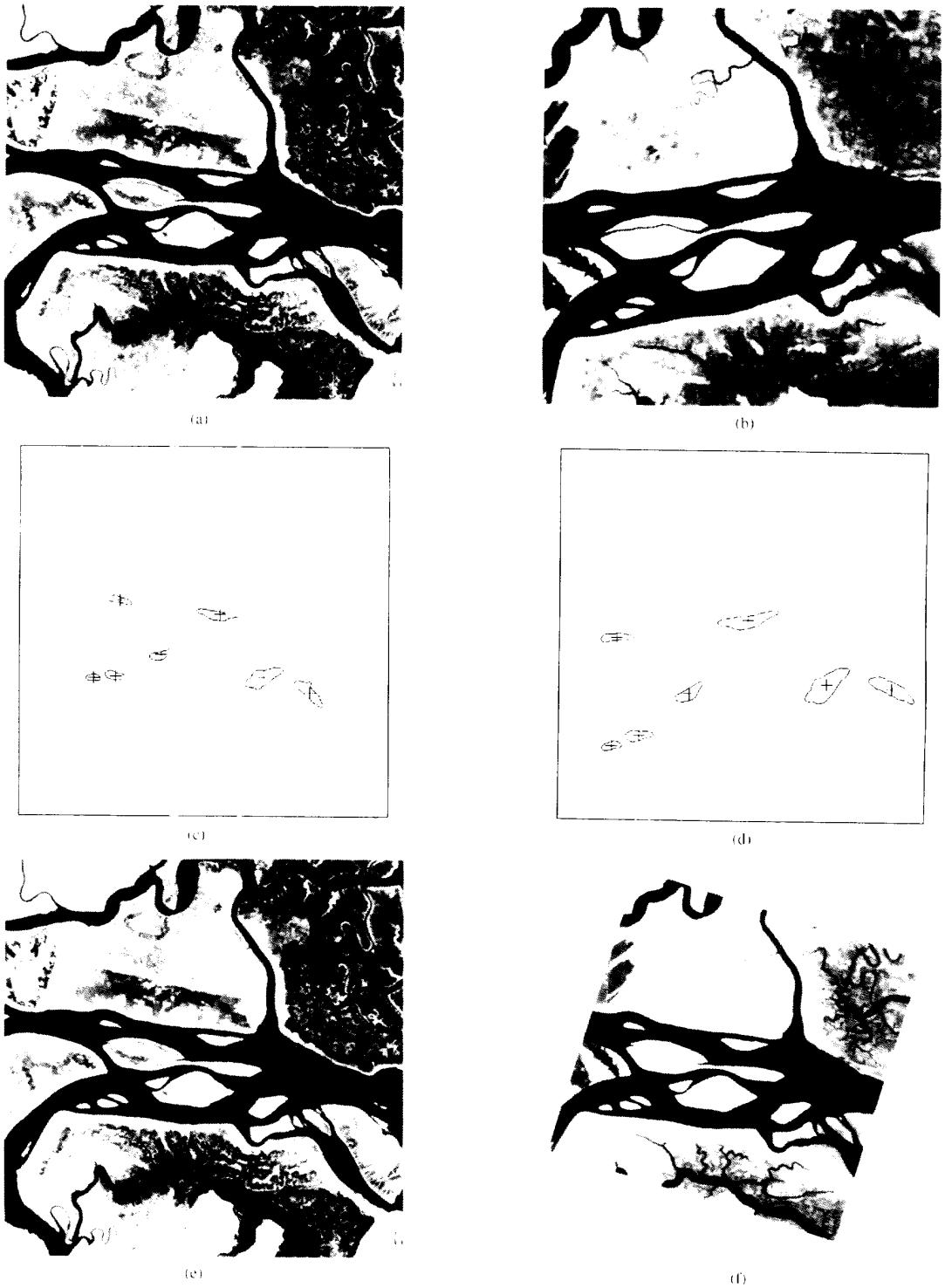

Fig. 7. (a), (b) Landsat image and a Spot image, respectively, with scale difference; (c), (d) matched closed contours; (e), (f) Landsat image and the matching Spot image after transformation.

data set after translations and rotation. In this way, the ground truth of the transformation parameters between the test images is known. The proposed algorithm performed successfully for the cases in which at least half of the image extents overlap. One example is shown in Fig. 5, where (a) and (b) correspond to the zeroth and eighth bands of

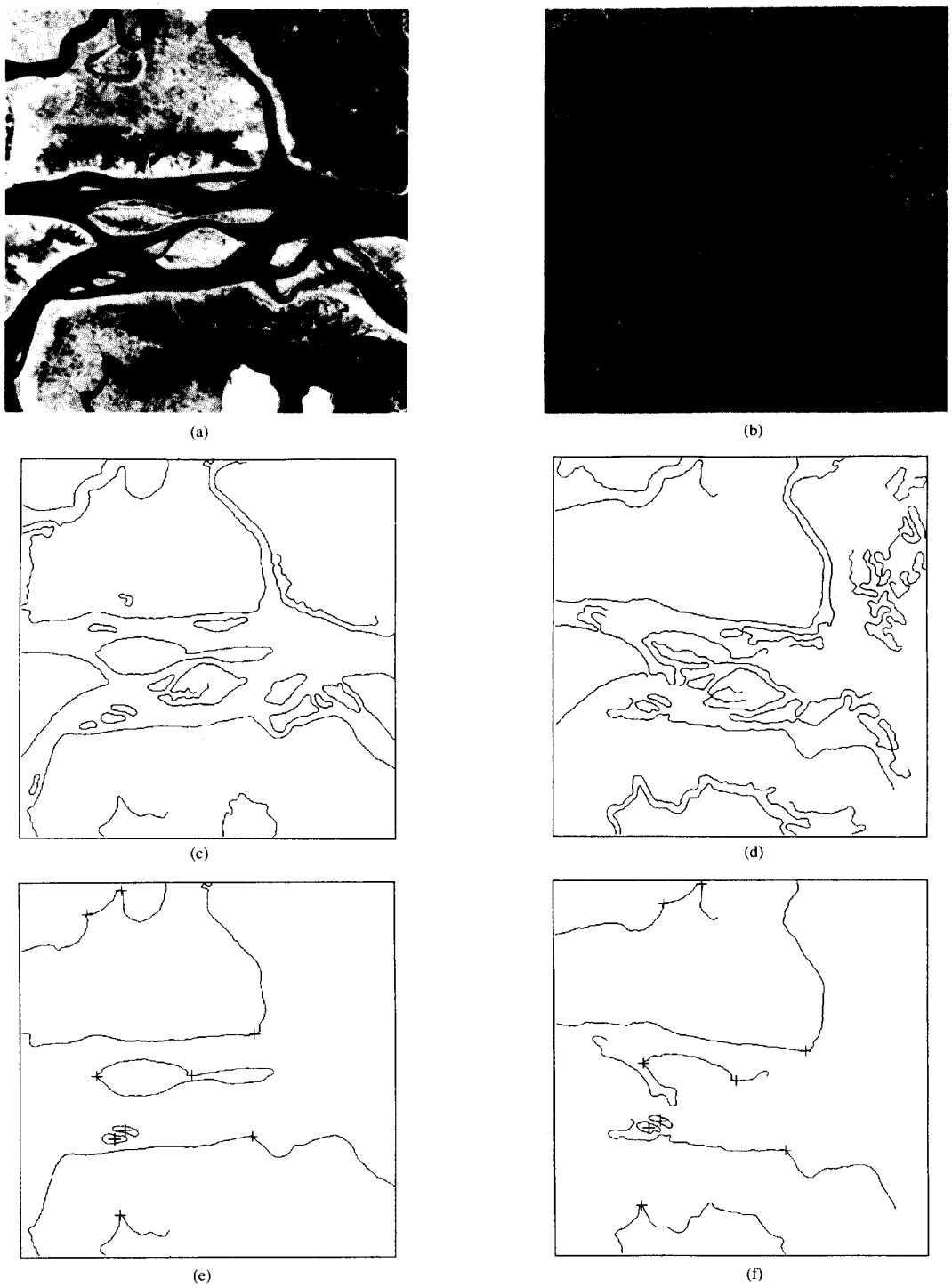

Fig. 8. (a), (b) Landsat image and a Seasat (SAR) image, respectively; (c) contour map of (a); (d) contour map of (b); (e), (f) correspondence between Landsat and Seasat contour maps in the first stage of the matching process.

the 12-band data set. Six pairs of closed contours and 9 pairs of salient segments have been matched, as shown in Fig. 5(c)–(d). The mosaic of the matched images is shown in Fig. 5(e). It can be seen from Table I that the estimated transformation parameters are very close to the ground truth.

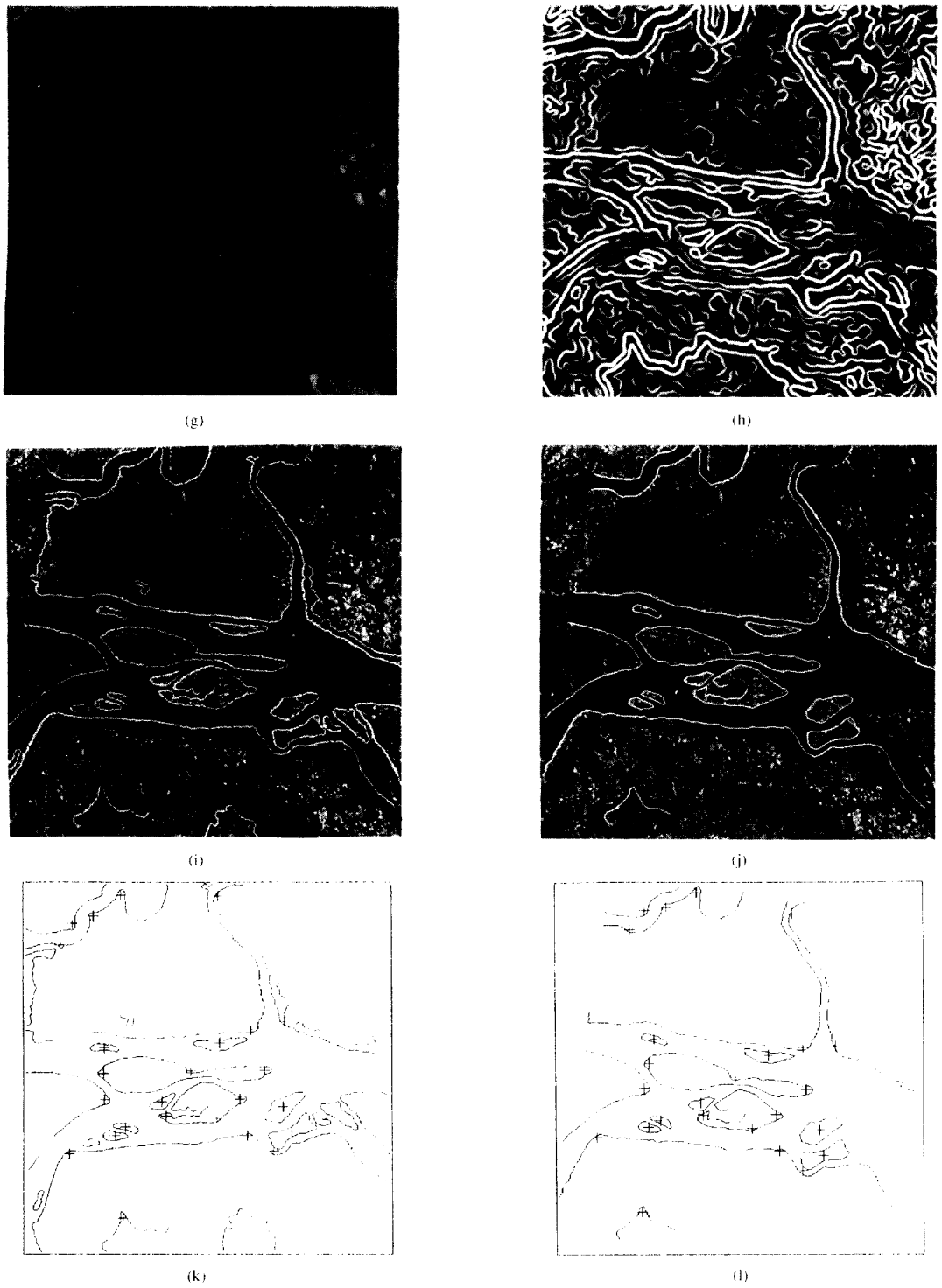

Fig. 8. *(continued)* (g) Smoothed SAR image. (h) image force used in the active contour model; (i) in the fine tuning stage, the Landsat contours are transformed and superimposed on the Seasat image as the initial condition; (j) converged contours of the active contour algorithm; (k), (l) final matching results, with the "+" marks denoting the centroids of the closed contours and the salient segments of the open contours.

Multisensor Optical-to-Optical Image Registration: Fig. 6(a)–(b) shows a pair of Landsat TM and Spot images. Twenty-eight matched points have been found, as shown in Fig. 6(c) (d). We rotated the Spot image by 90°, and applied

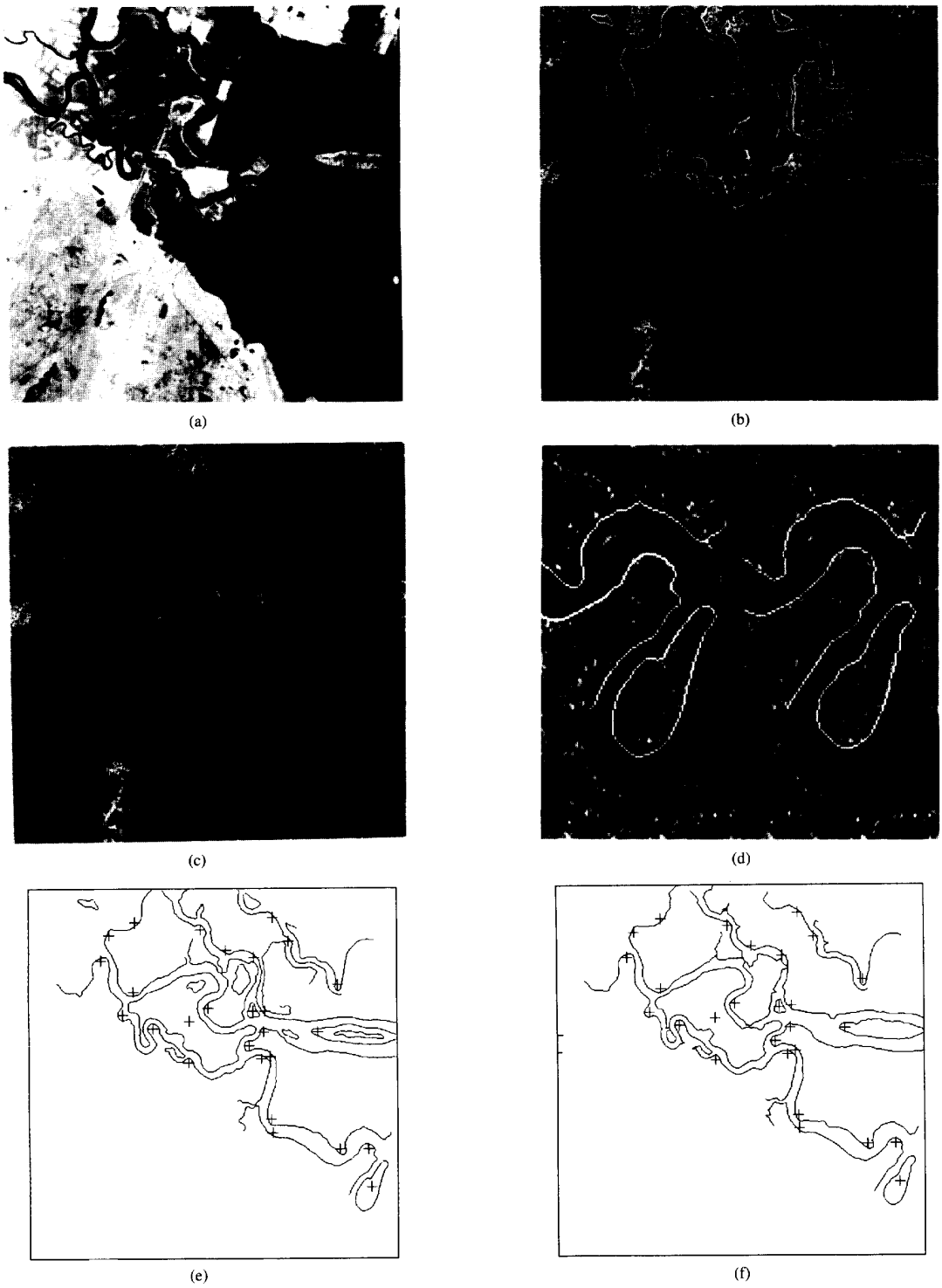

Fig. 9. (a), (b) Spot image and a Seasat (SAR) image, respectively—in (b), the contours from the Spot image are superimposed as the initial condition for the active contour algorithm; (c) converged contours; (d) blow-ups of three contours before and after convergence; (e), (f): matching results, with the "+" marks denoting the centroids of the closed contours and the salient segments of the open contours.

our method again. Twenty-seven matched points have been found in this case and the matched contours are shown in Fig. 6(e)-(f). Fig. 7(a)–(b) shows the same Landsat image and the scaled-up version of the Spot image. Seven pairs of

matched closed contours are identified in Fig. 7(c)–(d). The matched images after transformation are in Fig. 7(e)–(f).

Multisensor Optical-to-SAR Image Registration: The first example involves a Landsat (optical) image and a Seasat (SAR) image. Both have been rectified and resampled to the same pixel spacing as shown in Figs. 8(a)–(b). There is considerable displacement between them, such that the active contour model cannot be directly applied. Fig. 8(c)–(d) show the edges extracted from the optical image and the despeckled SAR image. It can be seen that the SAR edge map contains many spurious contours. The result of initial registration using the basic contour matching scheme is shown in Fig. 8(e)–(f) in which nine control points have been found. Fig. 8(g)–(i) show the smoothed SAR image (geometric filtering, 8eight iterations) and the image force for the active contour model, respectively. In the fine tuning stage, a total of 19 contours from the Landsat image are transformed and placed over the SAR image, as shown in Fig. 8(i). These contours are next used as the initial condition for the active contour algorithm. Due to the difference between the resolutions of two sensors, not all initial contours from the optical image have corresponding region boundaries in the SAR image. In this example, 12 converged contours are accepted after a thresholding test with respect to the averaged image force, and these contours are shown in Fig. 8(j). Fig. 8(k)–(l) show the final matching result of 25 pairs of closed contours and salient segments, with "+" denoting centroid positions.

The next example is concerned with the registration of Spot and Seasat images. These two images have been rectified and geocoded to within 5 pixels, and as a result the active contour algorithm is directly applicable. Fig. 9(a) shows the Spot image and Fig. 9(b) shows the Seasat image with the contours from the Spot image directly superimposed. The distances from the initial contours to the correct edge locations in the SAR image range from 1 to 5 pixels. Fig. 9(c) shows the converged contours. A blow-up of three initial and converged contours in Fig. 9(d) illustrates the active contour model's ability to match deformed region boundaries. Twenty-seven matched points have been obtained and are shown in Fig. 9(e)–(f).

It can be seen from Table I that the estimations of four transformation parameters produced by manual and automatic methods are very close. In each case, the proposed algorithms achieve a lower RMSE than manual registration. In general, optical-to-SAR image registration, either manual or using computerized algorithms, results in a higher RMSE due to the noise in the SAR images. We have also tested our algorithms on other data sets, including image pairs that do not cover the same region. The basic contour-matching algorithm is quite robust and reliable as long as corresponding contours are available. The registration scheme would fail if insufficient contour information can be extracted. In all cases in which the input images did not contain corresponding regions, matched points were not found, as expected.

V. Discussion

In this paper, we have presented two contour matching schemes for image registration. In the first scheme the contours are detected separately from two images and are matched at a later stage. A chain-code-based matching algorithm and an iterative consistency checking scheme have been developed. For single-sensor and multisensor optical image registration, the first algorithm is quite robust and reliable when the corresponding contours are available. It can handle images with large rotation and translation, and images with scale difference if there exist sufficient numbers of closed regions. A second scheme is developed for optical-to-SAR image registration. Assuming that multisensor images can be roughly aligned by rectification and geocoding, an elastic matching scheme based on the active contour model is proposed. The motivation here is to tackle the following two problems: (a) The difficulty in extracting well-defined, unbroken contours from a noisy SAR image; and (b) the deformation between images taken at different times by different sensors (e.g., river boundaries can change significantly after seasonal flooding). We found that the image force of the active contour model had to be computed based on the smoothed SAR image in order for contours to converge correctly. The filtering inevitably affects the accuracy of the contour locations. The active contour model also tends to smooth the sharp corners. Nevertheless, the registration error incurred by blurred contour locations is compensated to some extent by using the centroids of closed contours and corner segments as control points, since small deviations from the true boundary locations along the contour may cancel each other. The limitation of this algorithm lies in the requirement that well-defined, strong contours must be detected from the optical image as initial contours.

Previous researchers have noted that the task of automated multisensor image registration is very complex [6]. A combination of multiple techniques within the framework of an artificial intelligence system may be necessary. The proposed contour-based algorithms can serve as an important building block for such a system.

Acknowledgment

The authors would like to thank E. Rignot of the Jet Propulsion Laboratory, California Institute of Technology, and J. Star of the Geography Department, University of California, Santa Barbara, for providing us with the multisensor and multispectral data used in this paper. We also thank the anonymous reviewers for their constructive comments and suggestions.

References

[1] J. Curlander and R. McDonugh, *Synthetic Aperture Radar: Systems and Signal Processing*. New York: Wiley, 1991.
[2] W. Pratt, *Digital Image Processing*. New York: Wiley, 1991.
[3] R. Haralick and L. Shapiro, *Computer and Robot Vision, Volume 2*. Reading, MA: Addison-Wesley, 1993.
[4] Q. Zheng and R. Chellappa, "A computational vision approach to image registration," *IEEE Trans. Image Processing*, pp. 311–326, July 1993.
[5] B. Manjunath *et al.*, "A robust method for detecting image features with application to face recognition and motion correspondence," in *Proc. Int. Conf. Pattern Recogn.* (The Hague, Netherlands), Sept. 1992, pp. 208–212.

[6] E. Rignot *et al.*, "Automated multisensor registration: Requirements and techniques," *Photogrammetric Eng., Remote Sensing*, vol. 57, pp. 1029–1038, Aug. 1991.
[7] A. Ventura, A. Rampini, and R. Schettini, "Image registration by recognition of corresponding structures," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 305–314, May 1990.
[8] A. Goshtasby, G. Stockman, and C. Page, "A region-based approach to digital image registration with subpixel accuracy," *IEEE Trans. Geosci. Remote Sensing*, vol. 24, pp. 390–399, May 1986.
[9] Y. Wu and H. Maître, "A multiresolution approach for registration of a spot image and a SAR image," in *Proc. Int. Geosci. Remote Sensing Symp.*, May 1990, pp. 635–638.
[10] J. Greenfeld, "An operator-based matching system," *Photogrammetric Eng., Remote Sensing*, vol. 57, pp. 1049–1055, Aug. 1991.
[11] T. Schenk, J. Li, and C. Toth, "Toward an autonomous system for orienting digital stereopairs," *Photogrammetric Eng., Remote Sensing*, vol. 57, pp. 1057–1064, Aug. 1991.
[12] S. Mason and K. Wong, "Image alignment by line triples," *Photogrammetric Eng., Remote Sensing*, vol. 58, pp. 1329–1334, Sept. 1992.
[13] N. Nasrabadi, "A stereo vision technique using curve-segments and relaxation matching," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 566–572, May 1992.
[14] K. Boyer and A. Kak, "Structural stereopsis for 3-D vision," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, pp. 144–166, Mar. 1988.
[15] R. Touzi, A. Lopes, and P. Bousquet, "A statistical and geometrical edge detector for SAR images," *IEEE Trans. Geosci., Remote Sensing*, vol. 26, pp. 764–772, Nov. 1988.
[16] J. Chen, A. Huertas, and G. Medioni, "Fast convolution with Laplacian-of-Gaussian masks," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, pp. 584–590, July 1987.
[17] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 679–698, Nov. 1986.
[18] R. Duda and P. Hart, *Pattern Classification and Scene Analysis*. Reading, MA: Addison-Wesley, 1973.
[19] A. Goshtasby, "Image registration by local approximation methods," *Image Vision Comput.*, vol. 6, pp. 255–261, Nov. 1988.
[20] L. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, pp. 325–376, Dec. 1992.
[21] A. C. Bovik, "On detecting edges in speckle imagery," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1618–1627, Oct. 1988.
[22] T. Grimmins, "Geometric filter for speckle reduction," *Appl. Optics*, vol. 24, pp. 1438–44, May 1985.
[23] A. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
[24] J. Curlander, R. Kwok, and S. Pang, "A post-processing system for automated rectification and registration of spaceborne sar imagery," *Int. J. Remote Sensing*, vol. 8, pp. 621–638, Apr. 1987.
[25] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," in *Proc. 1st Int. Conf. Comput. Vision*, 1987.
[26] D. Williams and M. Shah, "A fast algorithm for active contours and curvature estimation," *CVGIP: Image Understanding*, vol. 55, pp. 14–26, Jan. 1992.
[27] A. Amini, S. Tehrani, and T. Weymouth, "Using dynamic programming for minimizing the energy of active contour in the presence of hard constraints," in *Proc. 2nd Int. Conf. Comput. Vision*, 1988.
[28] A. Cideciyan *et al.*, "Registration of high resolution images of the retina," in *Proc. SPIE*, vol. 1652, pp. 310–322, 1992.

Hui Li (S'90) was born in Xiamen, China, in 1967. He received the B.S. degree in electrical engineering in 1989 from the Fudan University, Shanghai, China, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of California at Santa Barbara in 1991 and 1993, respectively.

After graduation, he joined the Space Systems Division, Rockwell International Corp., Downey, CA. His research interests include image enhancement, image matching, data fusion, and image analysis.

B. S. Manjunath (M'91) received the B.E. degree in electronics from Bangalore University in 1985, the M.E. degree in systems science and automation from the Indian Institute of Science in January 1987, and the Ph.D. degree in electrical engineering from the University of Southern California (USC) in 1991.

From September 1987 to July 1991, he was a research assistant at the Signal and Image Processing Institute at USC, and in the summer of 1990, he worked at the IBM T. J. Watson Research Center, Yorktown Heights, NY. He joined the Electrical and Computer Engineering Department at the University of California, Santa Barbara, as an assistant professor in July 1991.

Dr. Manjunath was awarded the Bangalore University Gold Medal for the best graduating student in electronics engineering in 1985. He was also a recipient of a National Merit Scholarship from the Government of India during the period 1978–1985. His current research interests include computer vision, medical image analysis, pattern recognition, and neural networks.

Sanjit K. Mitra (SM'69–F'74) received the B.Sc. (Hon.) degree in physics from Utkal University, Cuttack, India; the M.Sc. (Tech.) degree in radio physics and electronics in 1956 from Calcutta University; and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1960 and 1962, respectively.

He is currently with the Center for Information Processing Research, Department of Electrical and Computer Engineering, University of California at Santa Barbara.

Dr. Mitra in May 1987 was awarded an Honorary Doctorate of Technology degree from the Tampere University of Technology, Finland.